sQMRA2

(swift Quantitative Microbiological Risk Assessment version 2)

MANUAL

JE Chardon, EG Evers

Centre for Zoonoses and Environmental Microbiology (Z&O),
Centre for Infectious Disease Control (CIb)
National Institute for Public Health and the Environment (RIVM),
Bilthoven, The Netherlands

March 2014

This investigation has been performed by order and for the account
of the Netherlands Food and Consumer Product Safety Authority (NVWA),
within the framework of project V/330371/12/QS.

CONTENTS

1 Introduction

Quantitative Microbiological Risk Assessment (QMRA) is a methodology to evaluate microbiological public health risks, which can be food-, direct contact-, or environmental-related. We focus on food here. In QMRA, mathematical modelling is usually applied to describe a food production chain. These models, together with available data, are used to calculate the presence and propagation of pathogens in a specific production chain, and the exposure of pathogens (the 'dose') to consumers. Dose-response models are subsequently used to estimate the number of cases of illness. Main uses of the sQMRA approach are the capability to compare the risk of pathogen-food product combinations, and to increase insight in the concepts of QMRA for people new in this working field.

1.1 sQMRA-tool (deterministic version)

Classical QMRA's are very time consuming due to complicated modelling and the collection of necessary data. To answer microbiological risk questions quicker, in 2006 a tool was developed based on a simplified deterministic modelling approach. It was called "sQMRA-tool" (swift Quantitative Microbiological Risk Assessment tool) and was developed in Microsoft Excel XP. (Evers & Chardon, 2010). Special attention was given to make the sQMRA tool insightful, for educational purposes. Like in full scale QMRA, pathogen numbers are followed through the food chain, which in this case starts at retail and ends with the number of human cases of illness. The first sQMRA-tool is deterministic and includes cross-contamination and preparation (heating) in the kitchen and a dose-response relationship. The general setup of the sQMRA tool consists of consecutive questions for values of each of the 11 parameters, always followed by intermediate model output broken down into categories of contamination, cross-contamination and preparation. In a separate sheet, model input and output are summarized and exposure as well as cases are attributed to the distinguished categories. As a relative risk measure, intermediate and final model outputs are always compared with results from a full scale QMRA of Campylobacter on chicken fillet.

1.2 sQMRA2

In 2014, a new, probabilistic, sQMRA-tool (named sQMRA2) is completed, based on the first deterministic version.

The sQMRA2 model is a risk assessment model in which the propagation of a pathogen is followed, starting at the retail phase. Processes considered are storage, cross-contamination and preparation (=heating, in the sense of cooking, frying, etc. of the product) in the kitchen, which leads to nineteen categories of portions. The resulting exposure is input for the effect modeling part, which expresses risk in terms of infection, illness, DALY and C.O.I., both at the portion and population level.

sQMRA2 is implemented as an @RISK/Excel spreadsheet and consists of a MODEL sheet, a RESULTS sheet and a REFERENCE DATA sheet. In the MODEL sheet, parameter values are inserted, and intermediate

calculation results are presented after every module, such as the number of portions per category and the number of cfu (colony forming units) per portion. In the RESULTS sheet, we present 1) a list of input parameter values, 2) attribution of exposure and of cases to the different transmission routes (storage and preparation) and 3) the relative risk (at portion and population level) at several intermediate points (no. of contaminated portions, no. of cfu) and at the end point (no. of human cases, DALY and C.O.I). sQMRA2 can be used deterministic as well as probabilistic (only variability is included, not uncertainty). A lot of attention was paid to design a user-friendly interface which makes sQMRA2 easy to use.

For sQMRA2 the palisade @RISK software (http://www.palisade.com) has to be installed on the computer.

The main differences of sQMRA2 in comparison with the sQMRA-tool (deterministic version) are:

MODEL sheet:

- Capability to execute probabilistic model calculations
- Variation in iteration values is presented by a central tendency measure and a lower and upper percentile. In addition, the result of individual iterations is presented in 'yellow' cells.
- Addition of storage at the consumers home prior to preparation as a module in the model
- Addition of portion categories on an aggregated level
- Cross-contamination module is simplified given the limited data availability.
- Extended heating module (D/z- model)
- Addition of a choice between two dose-response models in the infection and illness module
- Addition of DALY (disability-adjusted life years) and C.O.I. (cost-of-illness) as health metrics. See also Chapter 3.3.6.

RESULTS-sheet

- Addition of attribution of exposure and effect at the storage level
- Addition of relative risk points of comparison
- Addition of relative risk on portion level
- Addition of multiple relative risk reference data
- Addition of the probability of exposure
- Addition of variability on portion level and statistical uncertainty in retail data

2 Model outline

In Figure 1, a schematic overview of the sQMRA2 model is shown. This will be useful to keep the 'big picture' in mind when studying the rest of this manual. The parameter symbols in Figure 1 will be explained in Chapter 3.

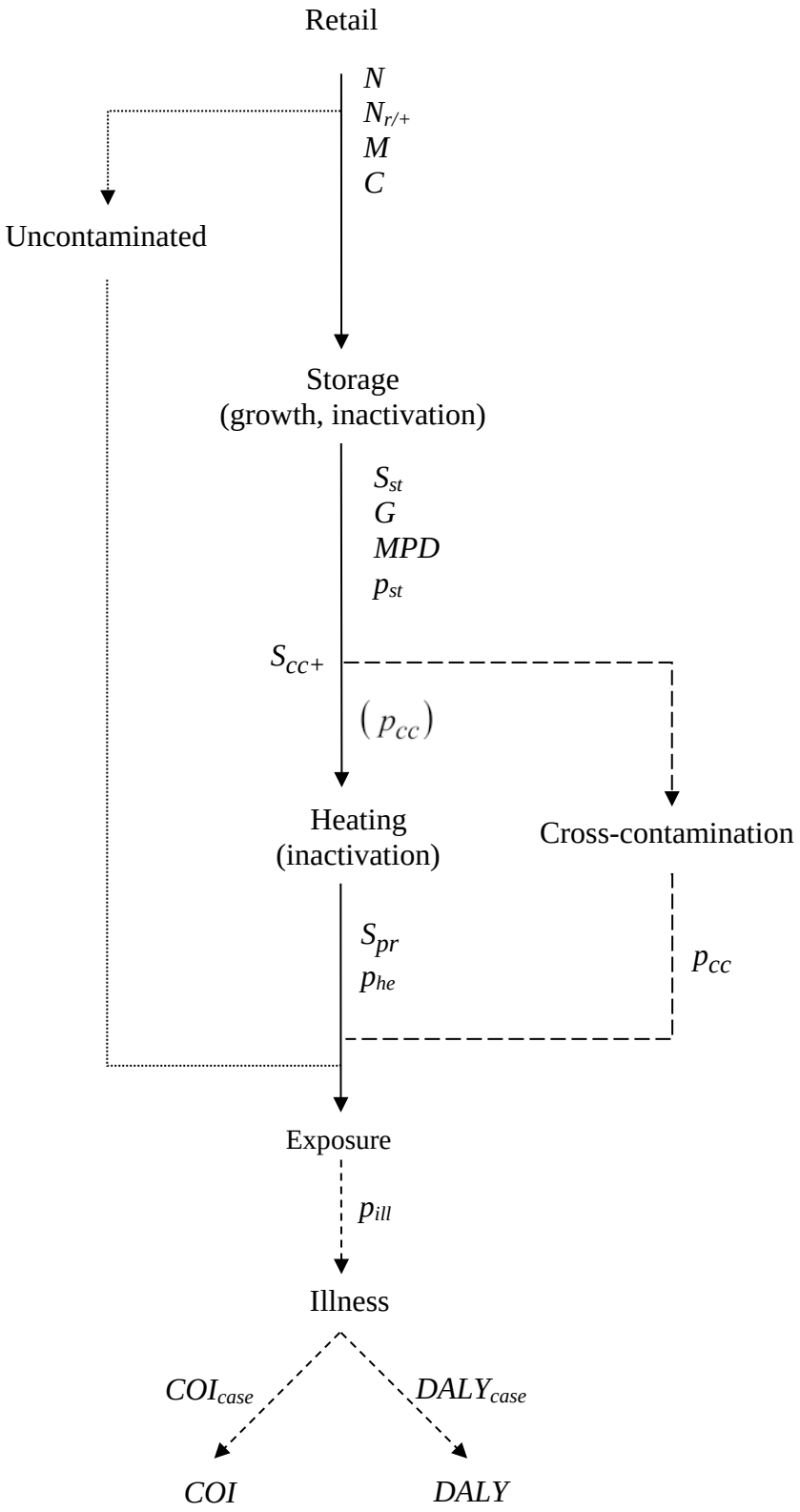


Fig.1: Overview of the sQMRA2 model.

3 MODEL sheet

On the MODEL sheet of sQMRA2, parameter values have to be entered and intermediate calculation results are shown.

3.1 Structure

The MODEL sheet consists of 5 modules with each the same visual structure. The five modules are: consumption and retail, storage, cross contamination, preparation and infection & illness. In Figure 2 all modules are shown.

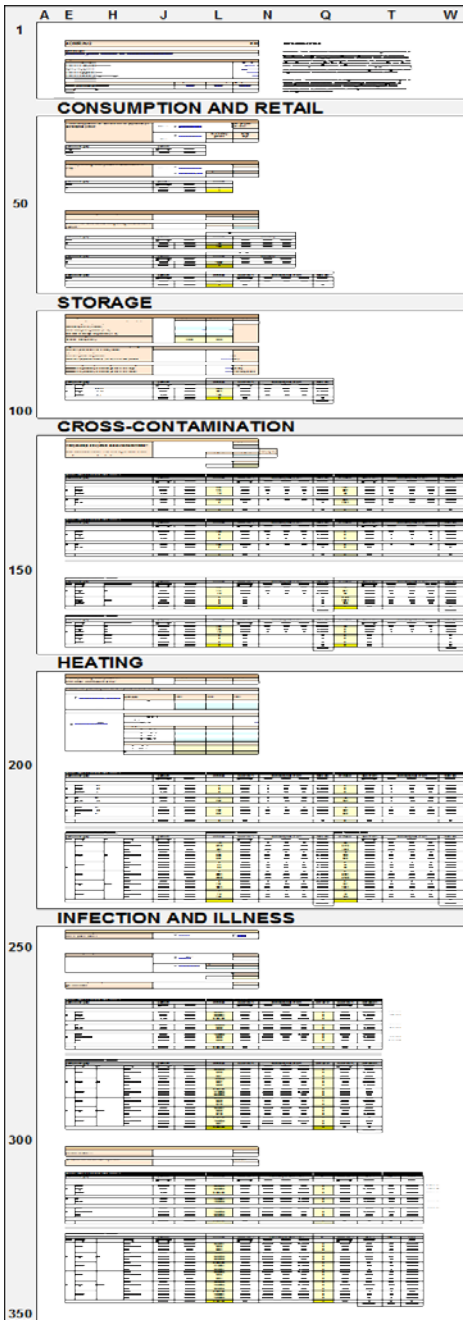


Fig. 2 : the 5 modules on the MODEL sheet

The modules contain the following visual characteristics:

- *Model input:* Peach colored questions blocks.
- *Model output:* Intermediate calculation results for the portion categories, for the individual portion categories (white table headers) as well as on an aggregated level (black table headers).

3.1.1 Model input

In figure 3, we show question 5 as an example of a peach colored input block.

	E	F	G	H	I	J	K	L	M
76									
77	**5. storage conditions**						room (%)	fridge (%)	freezer (%)
78	What % of the portions is stored in each storage category ?						5.0%	80.0%	15.0%
79	Mean storage time (in hours)						10	100	
80	Max storage time (in hours)						100	1000	
81	Mean storage temperature (in ºC)						20	5	
82	Std. Dev. of storage temperature (in ºC)						2	1	
83	Iteration: storage time						0.95	120.87	
84	Iteration: storage temp						18.86	5.73	

Fig. 3: example of input block.

All data entry cells have a blue font color. To run a simulation with deterministic parameter values, it is sufficient to enter data in the white colored data entry cells. If a simulation including variability is favored, then data also have to be entered in the corresponding green data entry cells. In this manual we include screenshots from a risk assessment example including variability in every module of the tool, to better demonstrate the variability of statistic model outputs. The default setting of the tool however is deterministic (no data in green data entry cells). In some question blocks, yellow iteration cells are included (e.g. cells K83-L84). They have an educational purpose as they show individual iteration values as used in the model calculations.

3.1.2 Model output

The following structure is used throughout the MODEL sheet, for individual and aggregated portion categories:

Individual portion categories

In figure 4 we display the calculation results from the storage module as a typical example of the individual portion categories.

	E	F	G	H	I	J	K	L	M	N	O	P	Q
96								CFU AFTER STORAGE					
97	**portion category**					**portions**		**iteration**	**non-zero's**	**variability non-zero's**			**total cfu**
98	retail	storage				percentage	number		percentage	2.50%	mean	97.50%	number
99		room				0.500%	6.1E+06	6667	99.8%	30	2.3E+07	2.8E+06	1.4E+14
100	+	fridge				8.0%	9.7E+07	7272	98.0%	5	1.6E+04	9.2E+04	1.5E+12
101		freezer				1.5%	1.8E+07	574	97.8%	2	1421	9326	2.5E+10
102	-					90.0%	1.1E+09	0	0%				0
103													1.4E+14

Fig. 4: calculation results from storage module

Notice the following items:

- *Portion category (cells E97-I102):* Indication of the portion categories.

- *Portions (cells J97-K102):* The division of portions in portion categories as percentage (cells J98-J102) and as number of portions (cells K98-K102).
- *The dimension of the calculation results (cell L96)*
- *Iteration (cells L97-L102):* The yellow colored iteration cells display individual iteration values. For instance, this can be number of cfu on a portion (e.g. 6667 cfu on the portion after storage at room temperature, cell L99), or probabilities of infection or illness. In all blocks of iteration cells within the individual portion categories, one cell is dark yellow. This dark yellow cell represents a randomly sampled portion throughout the whole calculation chain, which is done for educational purposes.
- *Non-zero's (cells M97-M102):* The percentage of non-zero's in the simulation. This can be the percentage of contaminated portions for this category at this point in the food chain (e.g. 99.8% of the portions of the category '+, room' is contaminated, cell M99), or iterations of probabilities of infection or illness higher then zero.
- *Variability non-zero's (cells N97-P101):* The statistics of the non-zero iteration values contain a lower percentile, a central tendency measure and an upper percentile.
- *total cfu (cells Q97-Q102):* For every portion category the total number of cfu in that portion category is given. E.g. in the category room temperature, there are $6.1 \cdot 10^6$ portions, 99.8% is still contaminated after storage with a mean of $2.3 \cdot 10^7$ cfu. Therefore the total amount of cfu (cell Q99) in this portion category is $6.1 \cdot 10^6 * 99.8\% * 2.3 \cdot 10^7 = 1.4 \cdot 10^{14}$ cfu.
- *total cfu for all portion categories (cell Q103)*: A summation of all cfu in all categories.

Error handling is implemented as follows: If the percentage non-zero's in a specific portion category is 0%, and therefore no portions in the category are contaminated, the variability of non-zero's is not applicable and will be greyed out by means of conditional formatting.

Aggregated portion categories

In figure 5, we display the aggregated portion categories for the block: 'cfu on portion after heating', as an example.

	E	F	J	K	L	M	N	O	P	Q
200	**AGGREGATED PORTION CATEGORIES**				**CFU ON PORTION AFTER HEATING**					
201	**portion category**		**portions**		**iteration**	**non-zero's**	**variability non-zero's**			**total cfu**
202			percentage	number		percentage	2.50%	mean	97.50%	number
203										
204		room	0.500%	6.1E+06	6	19.9%	1	1.4E+07	5.9E+05	1.7E+13
205	+	fridge	8.0%	9.7E+07	1372	18.1%	1	7729	4.5E+04	1.4E+11
206		freezer	1.5%	1.8E+07	0	16.1%	0	762	5450	2.2E+09
207										
208	+	cc	5.0%	6.1E+07	1337	17.9%	1	7.6E+05	5.5E+04	8.3E+12
209		no cc	5.0%	6.1E+07	1372	17.9%	1	7.9E+05	4.3E+04	8.6E+12
210										
211		done	8.0%	9.7E+07	2	3.5%	1	1.6E+05	1.9E+04	5.4E+11
212	+	undercooked	1.5%	1.8E+07	1337	67.8%	1	7.5E+05	3.4E+04	9.2E+12
213		raw	0.500%	6.1E+06	7272	98.0%	4	1.2E+06	9.8E+04	7.1E+12
214										
215	-		90.0%	1.1E+09	0	0%				0

Fig. 5: aggregated portion categories

The column structure in blocks with aggregated portion categories is identical to the blocks with individual portion categories. The rows, by contrast, represent now data from *multiple* individual portion categories. So cells E204:Q204 -aggregated portion category "room"- now shows statistics about all portions that have been stored at room temperature, without looking at cross-contamination status or heating method. And, as a second example, the number of portions in the aggregated portion category "cc"($6.1 \cdot 10^7$ in cell K208) is a summation of the nine individual cross-contamination portion categories. For a technical description of the aggregation method we refer to Ch. 6, section: 'Using RiskDiscrete for aggregated portion categories'.

Error handling for the aggregated portion categories is complicated due to nested excel-formulas used to calculate the output. In a specific aggregated portion category, the whole portion category is greyed out by means of conditional formatting when:

- The percentage of portions in a specific portion category is 0%.
- The percentage non-zero's in a specific portion category is 0%.

3.2 Symbols

Most used symbols for MODEL sheet formulas (see Ch. 3.3) are *N* for number of portions, *S* for subdivision of these portions into fractions of categories of portions, *d* for the dose of cfu on a portion, *t* for time, *T* for temperature and *p* for probability.
Figure 6 (which is taken from the RESULTS sheet, see Ch. 4) gives an overview of the parameters used in the calculations.

Fig. 6: Model parameters

INPUT PARAMETERS

scope	
Pathogen	pathogen X
Food product	product Y
Population specification	population Z
Population size (POP)	1,0E+08
Consumption period in days (t_{cons})	365

question	symbol	value	unit
1 Portions consumed			
Point estimation: portions per person per month	N_{pppm}	not selected	/pppm
Foodsurvey data: total portions consumed in survey	N_{survey}	10	portions
Foodsurvey data: total number of surveydays	t_{survey}	300	d
2 Pathogen prevalence in retail			
Point estimation of prevalence	Cont	not selected	---
Surveillance data: samplesize	Size	1000	samples
Surveillance data: number of positive samples	Pos	100	samples
3 Portionsize			
Portionsize: mean	M_{mean}	100	g
Portionsize: st dev	M_{stdev}	10	g
4 Pathogen concentration			
Concentration: mean	$\log_{10}C_{mean}$	1,00	log10 (cfu/g)
Concentration: st dev	$\log_{10}C_{stdev}$	1,00	log10 (cfu/g)
5 Storage conditions			
% of portions stored in room	S_{st_room}	5,0%	---
% of portions stored in the fridge	S_{st_fridge}	80,0%	---
% of portions stored in freezer	$S_{st\_freezer}$	15,0%	---
Storage time room: mean	$t_{st\_mean\_room}$	10	h
Storage time room: maximum	$t_{st\_max\_room}$	100	h
Storage time fridge: mean	$t_{st\_mean\_fridge}$	100	h
Storage time fridge: maximum	$t_{st\_max\_fridge}$	1000	h
Storage temperature room: mean	$T_{st\_mean\_room}$	20	°C
Storage temperature room: st dev	$T_{st\_stdev\_room}$	2	°C
Storage temperature fridge: mean	$T_{st\_mean\_fridge}$	5	°C
Storage temperature fridge: st dev	$T_{st\_stdev\_fridge}$	1	°C
6 Growth and inactivation characteristics of pathogen			
Minimum generation time in foodproduct	t_{gen_min}	1	h
Optimum growth temperature	T_{opt}	37	°C
Minimum growth temperature	T_{min}	4	°C
Maximum population density	MPD	1,0E+09	cfu/g
Probability of survival in room	$p_{st\_day\_room}$	1,00	/d
Probability of survival in fridge	$p_{st\_day\_fridge}$	0,50	/d
Probability of survival in freezer	$p_{st\_freezer}$	0,10	---

INPUT PARAMETERS

Question	symbol	value	unit
7 Cross-contamination parameters			
% of portions possibly causing cross-contamination	S_{cc+}	50,0%	---
Fraction of CFU ingested: mean	$\log_{10}p_{cc\_mean}$	-3	log10
Fraction of CFU ingested: st dev	$\log_{10}p_{cc\_stdev}$	1,00	log10
Maximum transfer rate	$\log_{10}MTR$	0	log10
8 Heating categories			
% of portions done	S_{pr_done}	80,0%	---
% of portions undercooked	S_{pr_unco}	15,0%	---
% of portions raw	S_{pr_raw}	5,0%	---
9 Probability of survival during heating			
Probability of survival done: most likely	Phe_ml_done	not selected	---
Probability of survival undercooked: most likely	Phe_ml_unco	not selected	---
Probability of survival raw: most likely	Phe_ml_raw	not selected	---
Probability of survival done: minimum	$p_{he\_min\_done}$	not selected	---
Probability of survival undercooked: minimum	$p_{he\_min\_unco}$	not selected	---
Probability of survival raw: minimum	$p_{he\_min\_raw}$	not selected	---
Probability of survival done: maximum	$p_{he\_max\_done}$	not selected	---
Probability of survival undercooked: maximum	$p_{he\_max\_unco}$	not selected	---
Probability of survival raw: maximum	$p_{he\_max\_raw}$	not selected	---
DRT-value at reference temperature	DRT_{ref}	0,11	min
z-value	z	12,3	°C
Reference Temperature	T_{ref}	90	°C
Heating time done: mean	$t_{he\_mean\_done}$	10	min
Heating time done: st dev	$t_{he\_stdev\_done}$	1	min
Heating time undercooked: mean	$t_{he\_mean\_unco}$	5	min
Heating time undercooked: st dev	$t_{he\_stdev\_unco}$	1	min
Heating time raw: mean	$t_{he\_mean\_raw}$	0	min
Heating time raw: st dev	$t_{he\_stdev\_raw}$	0	min
Heating temperature done: mean	$T_{he\_mean\_done}$	90	°C
Heating temperature done: st dev	$T_{he\_stdev\_done}$	10	°C
Heating temperature undercooked: mean	$T_{he\_mean\_unco}$	70	°C
Heating temperature undercooked: st dev	$T_{he\_stdev\_unco}$	10	°C
Heating temperature raw: mean	$T_{he\_mean\_raw}$	18	°C
Heating temperature raw: st dev	$T_{he\_stdev\_raw}$	0	°C
10 Endpoint dose-response model	no parameter	Infection	---
11 Dose-response parameters			
Binomial parameter: r	r	not selected	---
Beta-binomial parameter: alpha	α	0,01	---
Beta-binomial parameter: beta	β	10,00	---
12 Probability of illness given infection	$p_{ill\|inf}$	0,10	---
13 DALY per case	$DALY_{case}$	1,0E-03	daly/case
14 Cost-of-illness per case	COI_{case}	100	euro /case

3.3 Executing a sQMRA

In Chapter 3.1, we described the structure of the MODEL sheet. In this chapter, we present all MODEL sheet questions. Furthermore, we give short user instructions when necessary and we show the mathematical equations underlying the sQMRA2 MODEL sheet. When in the MODEL sheet radio buttons are used for input selection, the name of the radio button is indicated before the concerning equation. Note, that in the tool user friendly percentages can be (and must be) entered for portion categories. All calculations are naturally done with fractions.

3.3.1 Scope

	E	F	G	H	I	J	K	L	M
9	**scope**								
10	What is the pathogen of interest?								pathogen X
11	What is the foodproduct of interest?								product Y
12	Specify the population								population Z
13	What is the population size?								1,0E+08
14	What is the consumption period in days?								365
15									
16							lower percentile	central tendency	upper percentile
17	Select preferred output statistics						2,50%	mean	97,50%
18	Enter preferred currency								euro

Characteristics of the risk assessment have to be entered in the scope (e.g. name of pathogen, food product, etc., see figure above). Only the data on population size *Pop* (cell M13) and consumption period t_{cons} (cell M14) is used in the model calculations. The remaining questions only have an administrative purpose.

Calculation results are presented throughout the tool. We implemented a variety of output statistics that can be selected in cells K17-M17. When one of these cells is selected, a dropdown list will appear where one of the following output statistics can be selected:

- Lower percentile: 25%, 10%, 5%, 2.5%, 1%, 0.1% and 0.01%
- Central tendency: mean, median or mode
- Upper percentile: 75%, 90%, 95%, 97.5%, 99%, 99.9%, and 99.99%

In Ch. 6, section ‘Calculation of statistics’, the technical implementation of the output statistics is discussed.
In cell M18, the currency used for the cost-of-illness calculations can be entered.

3.3.2 Consumption and retail

Question 1: Portions consumed

	E	F	G	H	I	J	K	L	M
23									
24	1. How many portions are consumed in the population per								port. per pers.
25	consumption period?					FALSE	○ point estimation		per month
26									1.00
27							◉ foodsurvey data	no. of survey	survey
28								portions	days
29						TRUE		10	300
30									
31	portion category					portions			
32	all portions					percentage	number		
33	+ and -					100.0%	1.2E+09		

Question 1 deals with the consumption rate. When the radio button is set to 'point estimation', data entered in 'number of portions per person per month' N_{pppm} (cell M26), is used to calculate the number of portions consumed N for the chosen population size and time period (cell K33):

Radio button 'point estimation': $$N = N_{pppm}\left(\frac{t_{cons}}{365/12}\right)Pop$$

If preferred and available, food survey data can be entered directly by selecting radio button 'food survey data'. Then the 'no. of survey portions' N_{survey}, (cell L29), which is the number of portions consumed during a specific food survey, and the number of survey days t_{survey} , (cell M29), which equals the number of survey participants multiplied with the length of the survey in days, must be entered. The number of portions consumed N then equals:

Radio button 'food survey data': $$N = \left(\frac{N_{survey}}{t_{survey}}\right)t_{cons}Pop$$

Question 2: Pathogen prevalence in retail

	E	F	G	H	I	J	K	L	M
36									
37	2. What percentage of the portions is contaminated at retail?								percent
38						FALSE	○ point estimation		10.0%
39							◉ surveillance data	sample size	positives
40						TRUE		1000	100
41									
42	portion category					portions		iteration	
43	retail					percentage	number		
44	+					10.0%	1.2E+08	1	
45	-					90.0%	1.1E+09	0	

Question 2 asks for the pathogen prevalence. When the radio button in question 2 is set to 'point estimation', data entered for 'the percentage of portions contaminated at retail' *Cont* (cell M38), is used to calculate the number of contaminated portions in retail $N_{r/+}$ (cell K44):

Radio button 'point estimation': $N_{r/+} = N \cdot Cont$

When the 'surveillance data' radio button is selected, surveillance data can be entered. The 'sample size' *Size* (cell L40) represents the number of samples taken, and the 'positives' *Pos* (cell M40) defines the number of positive samples detected:

Radio button 'surveillance data': $$N_{r/+} = N\left(\frac{Pos}{Size}\right)$$

The dark yellow iteration cell (cell L44) represents the random selection of one portion (hence "1" as value in this cell), see also Ch. 3.1.2. The statistical uncertainty of the number of portions consumed (question 1) and of the pathogen prevalence (question 2) is presented on the RESULTS sheet (see Chapter 4.6.2).

Questions 3 & 4: Portion size & Pathogen concentration

	E	F	G	H	I	J	K	L	M	N	O	P	Q
48													
49	**3. What is the size in grams of one portion?**							mean	st.dev				
50								100	10				
51													
52	**4. What is the concentration in log10 cfu/g in contaminated portions?**							mean	st.dev				
53								1.00	1.00				
54													
55								**PORTION SIZE (g)**					
56	**portion category**					**portions**		**iteration**	**variability**				
57	retail					percentage	number		2.50%	mean	97.50%		
58	**+**					10.0%	1.2E+08	101	81.35	100.00	120.53		
59	**-**					90.0%	1.1E+09	100	81.36	100.00	120.53		
60													
61								**CFU CONCENTRATION (cfu/g)**					
62	**portion category**					**portions**		**iteration**	**variability**				
63	retail					percentage	number		2.50%	mean	97.50%		
64	**+**					10.0%	1.2E+08	1.20	0.11	139.93	910.45		
65	**-**					90.0%	1.1E+09	0					
66													
67								**CFU ON PORTION**					
68	**portion category**					**portions**		**iteration**	**non-zero's**	**variability non-zero's**			**total cfu**
69	retail					percentage	number		percentage	2.50%	mean	97.50%	number
70	**+**					10.0%	1.2E+08	109	99.8%	11	1.4E+04	9.1E+04	1.7E+12
71	**-**					90.0%	1.1E+09	0	0%				0

Question 3: Portion size

Data entered for the mean portion size M_{mean} (cell L50) gives the portion size *M* (cell L58-59):

Deterministic: $M = M_{mean}$

When a standard deviation of the mean portion size M_{stdev} is entered (cell M50), the portion size *M* (cell L58-59) equals:

Variability: $M \sim Gamma((M_{mean}/M_{stdev})^2 , M_{stdev}^2/M_{mean})$

The unit ‘portion’ is used at this phase, although it can be a virtual concept when sale size in retail differs from portion size at the moment of consumption.

Question 4: Pathogen concentration

Data entered for the mean pathogen concentration $log_{10}C_{mean}$ (in $\log_{10}$ cfu/g, cell L53) equals pathogen concentration C (in cfu/g, cell L64).

Deterministic: $C = 10^{\log_{10} C_{mean}}$

In case of variability, C is assumed to be lognormally distributed. So when a standard deviation $log_{10}C_{stdev}$ of the mean concentration (also in $\log_{10}$ cfu/g, cell M53) is entered, C (in cfu/g, cell L64) equals:

Variability: $C \sim 10^{Normal(\log_{10} C_{mean}, \log_{10} C_{stdev})}$

Take care in the interpretation of the mean ($\mu = log_{10}C_{mean}$) and the st dev ($\sigma = log_{10}C_{stdev}$) of the concentration, as they are defined at log10-scale. So e.g. whereas the mean of $10^{Normal(1,0)}$ is 10 cfu/g, the mean of for example $10^{Normal(1,1)}$ is 142 cfu/g. For exact calculation of the mean concentration in cfu/g (*mean*), the following formula applies: $mean(10^{Normal(\mu,\sigma)}) = 10^{\left(\mu + \frac{\sigma^2 \ln 10}{2}\right)}$.

The exact formula for the standard deviation in cfu/g (*stdev*):

$stdev(10^{Normal(\mu,\sigma)}) = \sqrt{10^{(2\mu+\sigma^2 \ln 10)}\left(10^{\sigma^2 \ln 10} - 1\right)}$

(See Vose (2008), p. 659-660).

Alternatively, if one has available *mean* and *stdev*, one can calculate $\mu = log_{10}C_{mean}$ (cell L53) and $\sigma = log_{10}C_{stdev}$ (cell M53) as follows:

$$\mu = \frac{1}{2}\log_{10}\left(\frac{mean^4}{stdev^2 + mean^2}\right)$$

$$\sigma = \sqrt{\frac{\log_{10}\left(\frac{stdev^2}{mean^2} + 1\right)}{\ln(10)}}$$

Cfu on portion

The product of portion size M with pathogen concentration C is considered as the mean of a Poisson distribution, from which a sample is taken to obtain the dose of cfu’s in a portion at retail level d_r (cell L70)

$d_r \sim Poisson(M \cdot C)$

3.3.3 Storage

Questions 5: Storage conditions

	E	F	G	H	I	J	K	L	M
76									
77	**5. storage conditions**						room (%)	fridge (%)	freezer (%)
78	What % of the portions is stored in each storage category ?						5.0%	80.0%	15.0%
79	Mean storage time (in hours)						10	100	
80	Max storage time (in hours)						100	1000	
81	Mean storage temperature (in ºC)						20	5	
82	Std. Dev. of storage temperature (in ºC)						2	1	
83	Iteration: storage time						0.83	60.50	
84	Iteration: storage temp						20.87	5.53	

Question 5 'storage conditions' describes the storage of a portion from the consumer behavior perspective: location, time and temperature are the relevant parameters. We consider three household storage locations in sQMRA v2: 'room' (S_{st_room}), 'fridge' (S_{st_fridge}), and 'freezer' ($S_{st\_freezer}$). Storage at room temperature and in the fridge is considered to be a function of time and temperature, whereas storage in the freezer is modeled independent of time and temperature. The fraction of portions stored in each location can be entered in cells K78-M78.

For storage at room temperature and in the fridge, data entered for mean storage time t_{st_mean} (cells K79-L79) gives storage time t_{st} (cells K83-L83):

Deterministic: $t_{st} = t_{st\_mean}$.

In case of variability, t_{st} is assumed to be exponentially distributed. A maximum storage time t_{st_max} is then entered (cells K80-L80) and the storage time t_{st} equals:

Variability: $t_{st} \sim \text{ExponTrunc}(t_{st\_mean}, t_{st\_max})$

To describe variability in storage time we use a truncated exponential distribution. When the maximum storage time is set at a relative low value, the simulated mean storage time will be shorter than the entered mean storage time. For instance, when the maximum storage time entered in cell K80 is 5 times the average storage time as entered in K79, the simulated mean storage time mean(t_{st}) will be 3% shorter than the entered mean storage time t_{st_mean}.

Furthermore, for storage at room temperature and in the fridge, data entered for the mean storage temperature T_{st_mean} (cells K81-L81) gives storage temperature T_{st} (cells K84-L84).

Deterministic: $T_{st} = T_{st\_mean}$

In case of variability, T_{st} is assumed to be normally distributed. A standard deviation T_{st_stdev} of the storage temperature is then entered and the storage temperature equals:

Variability: $T_{st} \sim \text{Normal}(T_{st\_mean}, T_{st\_stdev})$

Question 6. Growth and inactivation characteristics of pathogen

	E	F	G	H	I	J	K	L	M
86	**6. growth and inactivation characteristics of pathogen**								
87	Minimum generation time in food product							1	hours
88	Optimum growth temperature							37	ºC
89	Minimum growth temperature							4	ºC
90	Maximum population density (MPD) in the food product?							1.0E+09	cfu/g
91									
92	What is the probability of survival per cfu at room temperature?							1.00	per day
93	What is the probability of survival per cfu in the fridge?							0.50	per day
94	What is the probability of survival per cfu in the freezer?							0.10	time independent
95									
96								**CFU AFTER STORAGE**	
97	**portion category**					**portions**		**iteration**	**non-zero's**
98	retail	storage				percentage	number		percentage
99		room				0.500%	6.1E+06	117	99.8%
100	+	fridge				8.0%	9.7E+07	118	98.0%
101		freezer				1.5%	1.8E+07	13	97.8%
102	-					90.0%	1.1E+09	0	0%

We assume that storage at room temperature or in the refrigerator results in either microbial growth or microbial inactivation, whereas storage in the freezer can only result in microbial inactivation. Microbial growth occurs when the storage temperature T_{st} is above the minimum growth temperature T_{min}. When storage temperature is equal to, or below the minimum growth temperature, microbial inactivation takes place. The growth and inactivation calculations are executed in cells G99-G101.

Growth

Four parameters determine the pathogen growth characteristics:

- t_{gen_min} (cell L87) is the minimum generation time for the pathogen in the selected food product.
- T_{opt} (cell L88) is the optimum growth temperature.
- T_{min} (cell L89) is the minimum growth temperature
- *MPD* (cell L90) is the maximum population density.

Generation times can be calculated from specific growth rates (time^{-1}) e.g. from literature using the equation: generation time = ln2/specific growth rate.

We calculate the dose of cfu's on a portion after storage, $d_{st|g}$ (cells L99-L101), by multiplication of the dose in retail d_r with a growth factor G and consider this to be the mean for a poisson distribution:

Storage given growth: $d_{st|g} \sim \text{Poisson}(d_r \cdot G)$

The growth factor G is a function of the storage time t_{st} of the portion and the generation time t_{gen} of the pathogen in an exponential primary growth model (Van Gerwen, 1998):

Growth: $G = \mathrm{e}^{(\ln 2/t_{gen})t_{st}}$

To calculate the influence of temperature on t_{gen}, we use the temperature part of the gamma model (van Gerwen, 1998) as secondary growth model:

Growth: $t_{gen} = t_{gen\_min} \Big/ \left(\frac{T_{st} - T_{min}}{T_{opt} - T_{min}} \right)^2$

Finally, to limit the growth of the pathogen in the portion, we use the maximum population density *MPD* (cfu/g):

Growth: If $d_{st|g} > MPD*M$ then $d_{st|g} = MPD*M$

Where M is the mass of the portion, (see question 3).

Inactivation

Inactivation is modelled as a stochastic process with a probability of survival per cfu, using a Binomial distribution. The dose on a portion after storage $d_{st|i}$ (cells L99-L101) is given by:

Storage given inactivation: $d_{st|i} \sim Binomial\ (d_r, p_{st})$

Where p_{st} is the probability of survival.

To describe inactivation for the locations room or fridge, the probability of survival of a cfu per day p_{st_day} is entered in cells L92-L93. The probability of survival p_{st} after a time period t (h) then equals:

Inactivation room & fridge: $p_{st} = \left(p_{st\_day}\right)^{(t_{st}/24)}$

For microbial inactivation in the freezer, no data on storage time is needed. The only parameter needed is the probability of a cfu to survive the storage in the freezer $p_{st\_freezer}$ (cell L94).

Inactivation freezer: $p_{st} = p_{st\_freezer}$

3.3.4 Cross-contamination

Question 7: Cross contamination parameters

	E	F	G	H	I	J	K	L	M	N
107										
108	**7. Cross-contamination parameters**								CC (%)	
109	What percentage of the portions leads to cross-contamination?								50,0%	
110										max. transfer
111	Given cross-contamination: What is the log10 fraction of cfu's							mean (log10)	st dev (log10)	rate (log10)
112	on a portion that ends up being ingested?							-3	1,00	0
113										
114									prob of cc/cfu	
115								iteration	2.8E-05	

The percentage of portions that leads to cross-contamination S_{cc+} can be entered in cell M109.

Calculating the transfer rate

The value entered for the mean transfer rate $\log_{10} p_{cc\_mean}$ (defined as the $\log_{10}$ fraction of cfu on a portion that ends up being ingested, cell L112) is converted to the cross contamination transfer rate p_{cc} (cell M115), as follows:

Deterministic: $p_{cc} = 10^{\log_{10} p_{cc\_mean}}$

In case of variability, p_{cc} is assumed to be lognormally distributed. The standard deviation $\log_{10} p_{cc\_stdev}$ of the mean transfer rate is then entered (also in $\log_{10}$ units, cell M112) and the transfer rate p_{cc} equals:

Variability: $p_{cc} \sim 10^{Normal(\log_{10} p_{cc\_mean}, \log_{10} p_{cc\_stdev})}$

To avoid unrealistic transfer rates (e.g. p_{cc} >1) we implemented an optional maximum transfer rate $\log_{10} MTR$ (also in $\log_{10}$ units, cell N112), which truncates p_{cc} to the desired maximum (default value of $\log_{10}$MTR is 0):

Variability: $p_{cc} \sim 10^{Normaltrunc(\log_{10} p_{cc_{mean}}, \log_{10} p_{cc_{stdev}}, \log_{10} MTR)}$

Modelling cross-contamination

In the sQMRA-tool, cross-contamination is modelled as two consecutive binomial processes. Given that cross-contamination occurs, as determined by S_{cc+}:

1. Every cfu on the portion has a probability $p_{cc}^{0.5}$ to transfer from the portion to the environment (knife, water tab, cutting board etc),
2. Every cfu in the environment has also a probability $p_{cc}^{0.5}$ to transfer from the environment to ingestion.

So it is assumed that the transfer from portion to environment and the transfer from environment to ingestion have the same probability.

	L	M	N	O	P	Q	R	S
147	**CFU TO ENVIRONMENT**						**CFU LEFT ON PORTION**	
148	**iteration**	**non-zero's**	**variability non-zero's**			**total cfu**	**iteration**	**non-zero's**
149		percentage	2.50%	mean	97.50%	number		percentage
150	5125	96.5%	1	7.4E+05	1.2E+05	2.2E+12	3.4E+04	99.8%
151	194	88.2%	1	1151	5395	4.9E+10	1212	97.9%
152	161	72.3%	1	136	636	8.9E+08	1149	97.6%
153	0	0%				0	3.9E+04	99.8%
154	0	0%				0	1406	98.0%
155	0	0%				0	1310	97.8%
156	0	0%				0	0	0%
157						2.2E+12		
158								
159	**CFU LEFT IN ENVIRONMENT**						**CFU TO INGESTION VIA CC**	
160	**iteration**	**non-zero's**	**variability non-zero's**			**total cfu**	**iteration**	**non-zero's**
161		percentage	2.50%	mean	97.50%	number		percentage
162	4458	96.4%	1	7.0E+05	1.1E+05	2.0E+12	667	70.7%
163	171	88.1%	1	920	4780	3.9E+10	23	53.5%
164	141	71.6%	1	109	544	7.1E+08	20	30.9%
165	0	0%				0	0	0%
166	0	0%				0	0	0%
167	0	0%				0	0	0%
168	0	0%				0	0	0%
169						2.1E+12		

Fig.7: Cross-contamination output

In Figure 7, a section of the individual portion categories block is presented, showing all doses related to cross-contamination.

Given that cross-contamination potentially occurs, the dose of cfu from the portion to the environment *denv* (cells L150-L152), equals:

$$denv \sim \text{Binomial}(d_{st}, p_{cc}^{0.5})$$

The remaining dose on the portion *dport* (cells R150-R152) is then described by:

$$dport = d_{st} - denv$$

The dose that will be ingested due to cross-contamination *dei* (cells R162-R164) is:

$$dei \sim \text{Binomial}(denv\ ,\ p_{cc}^{0.5})$$

And the dose that remains in the environment *del* (cells L162-L164) is:

$$del = denv - dei$$

Dose *dei* will be added to the dose on the portion after preparation, see Chapter 3.3.6, Infection and Illness.

In case of no potential cross-contamination,

denv (cells L153-155) = *dei* (cells R165-167) = *del* (cells L165-167) = 0

and

dport (cells R153-155) = d_{st}

3.3.5 Heating

Questions 8 & 9: Preparation categories & Probability of survival during heating

	E	F	G	H	I	J	K	L	M
173									
174	**8. What percentage of the portions is**						done (%)	undercooked (%)	raw (%)
175	**heated done, undercooked or raw?**						**80,0%**	**15,0%**	**5,0%**
176									
177	**9. What is the probability for a cfu on a portion to survive heating?**								
178									
179							done	undercooked	raw
180					estimators		most likely	most likely	most likely
181		○ direct estimation of survival			**mean or most likely**		1,0E-03	0,10	1,00
182					**minimum**		1,0E-04	0,01	1,00
183					**maximum**		0,01	0,50	1,00
184							0,004400001	0,202038001	1
185					inactivation characteristics				
186					DRT value at ref. temp. (min)				0,11
187					z-value (°C)				12,3
188					reference temp (°C)				90
189					heating parameters		done	undercooked	raw
190					**mean heating time (min)**		10	5	0
191		◉ DRT/z-inactivation model			**st dev heating time (min)**		1	1	0
192					**mean heating temp (°C)**		90	70	18
193					**st dev heating temp (°C)**		10	10	0
194					iterations		done	undercooked	raw
195					**heating time (min)**		8,53	5,24	0,0E+00
196					**heating temp (°C)**		89,44	68,06	18,00
197					**prob. of surv./cfu**		1,5E-70	0,16	1,00

Question 8: Preparation categories

The preparation module consists of three analogously modelled preparation categories. If desired, the categories can be used freely. For instance, one can perform a scenario analysis by setting all preparation parameter values to identical values except for one parameter.

We consider three preparation categories in sQMRA2: done (S_{pr_done}, cell K175), undercooked (S_{pr_unco}, cell L175) and raw (S_{pr_raw}, cell M175).

Question 9: Probability of survival during preparation

Cfu on a portion d_{port} have a certain probability to survive heating p_{he} (cells K197-M197).

When entering data to calculate the probability of survival p_{he} for a cfu on a portion, two options are offered: 'direct estimation of survival' and the 'D/z-inactivation model'.

Direct estimation of survival

Values entered for the mean probability of survival p_{he_mean} in cells K181-M181 give the probability of survival p_{he}.

Radio button 'direct estimation of survival' (deterministic): $p_{he} = p_{he\_mean}$

In case of variability, p_{he} is assumed to be Pert distributed. Then values for "minimum" (p_{he_min} cells K182-M182) and "maximum" (p_{he_max} cells K183-M183) parameters are entered, and cells K181-M181 are interpreted by sQMRA2 as "most likely" p_{he_ml}:

Radio button 'direct estimation of survival' (variability): $p_{he} \sim$ Pert (p_{he_min}, p_{he_ml}, p_{he_max})

DRT/z-model

In microbiology, an often used thermal inactivation model is the *D/z*-model. (van Asselt, 2006) which is termed here *DRT/z*-model as the symbol *D* already has a different meaning. The equation is:

DRT/z-model: $$p_{he} = 10^{-\frac{t_{he}}{DRT}}$$

Where *D* is the decimal reduction time (min):

$$DRT = DRT_{ref} \cdot 10^{-\left(\frac{T_{he} - T_{ref}}{z}\right)}$$

This model contains three parameters that describe the *inactivation characteristics* of a pathogen:

- DRT_{ref} (cell M186), the decimal reduction time expressed in minutes at temperature T_{ref}.
- z (cell M187), the temperature increase needed (in °C) to reduce the *D* value with a factor of 10.
- T_{ref} (cell M188), the reference temperature (in °C) at which DRT_{ref} was estimated.

And two *heating parameters* that describe the preparation by the consumer:

- T_{he}, the heating temperature (in °C) to which the portion is exposed.
- t_{he}, the heating time of the portion (in min).

These are explained further below.

DRT/z-model: heating parameters

The value entered for the mean heating time t_{he_mean} (cells K190-M190) gives the heating time t_{he} (cells K195-M195):

Radio button 'DRT/z-model' (deterministic): $t_{he} = t_{he\_mean}$

In case of variability, t_{he} is assumed to be normally distributed. Then also a standard deviation t_{he_stdev} (cells K191-M191) of the heating time is entered, and t_{he} equals:

Radio button 'DRT/z-model' (variability): $t_{he} = Maximum(0, \sim Normal(t_{he\_mean}, t_{he\_stdev}))$

The value entered for the mean heating temperature T_{he_mean} (cells K192-M192) gives the heating temperature T_{he} (cells K196-M196):

Radio button 'DRT/z-model' (deterministic): $T_{he} = T_{he\_mean}$

In case of variability, T_{he} is assumed to be normally distributed. Then also a standard deviation T_{he_stdev} (cells K193-M193) of the temperature is entered, and T_{he} equals:

Radio button 'DRT/z-model' (variability): $T_{he} \sim Normal(T_{he\_mean}, T_{he\_stdev})$

The dose on a portion after heating *dhe* (cells L222-L240, see figure 8) is given by:

$dhe \sim \text{Binomial}(dport, p_{he})$

	L	M	N	O	P	Q	R	S
219	CFU ON PORTION AFTER HEATING						TOTAL CFU INGESTED	
220	iteration	non-zero's	variability non-zero's			total cfu	iteration	non-zero's
221		percentage	2.50%	mean	97.50%	number		percentage
222	0	4.7%	1	2.3E+06	3.5E+05	2.6E+11	1.4E+04	71.5%
223	5.1E+05	74.3%	1	1.3E+07	6.5E+05	4.5E+12	5.3E+05	91.0%
224	5.3E+05	99.8%	28	2.3E+07	2.5E+06	3.4E+12	5.5E+05	99.8%
225	0	3.6%	1	1210	1.1E+04	1.7E+09	117	54.7%
226	4428	68.5%	1	5887	3.0E+04	2.9E+10	4545	82.6%
227	4583	97.9%	5	1.5E+04	8.5E+04	3.5E+10	4700	98.0%
228	0	2.6%	1	153	1550	2.9E+07	10	32.2%
229	295	61.3%	1	553	3569	4.6E+08	305	71.1%
230	307	97.6%	2	1324	8636	5.9E+08	317	97.7%
231	0	4.8%	1	2.4E+06	3.8E+05	2.7E+11	0	4.8%
232	6.1E+05	74.4%	1	1.4E+07	7.0E+05	4.7E+12	6.1E+05	74.4%
233	6.3E+05	99.8%	30	2.3E+07	2.8E+06	3.6E+12	6.3E+05	99.8%
234	0	3.6%	1	1276	1.2E+04	1.8E+09	0	3.6%
235	5194	68.7%	1	6271	3.1E+04	3.1E+10	5194	68.7%
236	5373	98.0%	5	1.6E+04	9.2E+04	3.7E+10	5373	98.0%
237	0	2.5%	1	169	1609	3.1E+07	0	2.5%
238	353	61.7%	1	590	3804	5.0E+08	353	61.7%
239	362	97.8%	2	1421	9326	6.3E+08	362	97.8%
240	0	0%				0	0	0%
241						1.7E+13		

Fig. 8: Cfu on portion after heating & total cfu ingested

Calculating the ingested dose.

The ingested dose d_{pr} (cells R222-R240) is calculated by adding the number of cfu ingested due to cross-contamination *dei* (cells R162-R168, see figure 7) to the number of cfu still on the portion after preparation *dhe*:

$$d_{pr} = dei + dhe$$

3.3.6 *Infection and Illness*

Questions 10 & 11: Outputs of the dose-response model & dose-response parameters

	E	F	G	H	I	J	K	L	M
245									
246	10. What is the output of the					INPUT:	◉ infection	INPUT:	○ illness
247	dose-response model?								
248									
249	11. Dose-response parameters								r
250							○ binomial	INPUT:	1.0E-03
251									
252								alpha	beta
253							◉ beta-binomial	INPUT:	0.01 10.00
254									
255									prob of inf. /cfu
256								iteration	0.01

Question 10. Outputs of the dose-response model

Both infection and illness are possible outputs of a dose response-model. In sQMRA2 one can select the desired output with a radiobutton. The default selection is 'infection'. When radio button 'illness' is selected, sQMRA2 sets the probability of illness given infection $p_{ill|inf}$ to the value 1 (see also question 12). In this case,

the intermediate calculation blocks for 'Probability of infection' are not relevant and hidden in the tool with a violet colour. This is indicated with blue text in sQMRA v2. The output of the dose-response model is termed p_{resp} (cell M256), which stands for the probability of infection (p_{inf}) or the probability of illness (p_{ill}) dependent on the chosen output in this question. Heading cell M255 is appropriately adjusted.

Question 11. Dose-response parameters

A choice is offered between the binomial and the beta-binomial dose-response model. By default, the binomial dose-response model is used in sQMRA2. Then a value for *r* (the probability of a response (infection or illness) with dose 1 cfu) must be entered in cell M250. The equation for the binomial dose-response model is:

radio button 'binomial' (deterministic): $$p_{resp} = 1-(1-r)^{d_{pr}}$$

If the radio button for the Beta-binomial dose-response model is selected, values for the α and the β parameters (cell L253 and cell M253) have to be entered for this model, which equals:

radio button 'beta-binomial' (variability): $$p_{resp} \sim 1-(1-\text{Beta}(\alpha,\beta))^{d_{pr}}$$

Question 12: Probability of illness given infection

	E	F	G	H	I	J	K	L	M
258	**12. What is the probability of illness**								Pill|inf
259	**given infection?**								0.10

When 'infection' is the output of the dose-response model (see question 10), the probability of illness given infection $p_{ill|inf}$ has to be entered in cell M259. The probability of illness p_{ill} is then calculated with:

$$p_{ill} = p_{inf}\, p_{ill|inf}$$

When 'illness' is selected as the output, Question 12 is hidden and a message is shown in blue text directing the user to Question 13.

Question 13 & 14: DALY per case & Cost-of-illness per case

	E	F	G	H	I	J	K	L	M
305									
306	**13. What is the amount of DALY**								DALY/case
307	**per case of illness?**								1,0E-03
308									
309	**14. What is the cost-of-illness per case?**								euro/case
310									100

In sQMRA2 we include the following health metrics (supplementary to cases of illness):

- DALY: The disability-adjusted life year (DALY) is a measure of overall disease burden, expressed as the number of years lost due to ill-health, disability or early death.
- C.O.I.: The cost-of-illness (C.O.I.) is a measure of the economic burden of disease.

The amount of DALY per case of illness $DALY_{case}$ (cell M307) and cost-of-illness per case COI_{case} (cell M310) that must be entered in Question 13 and 14 are used to calculate the number of DALYs and the COI per portion category as endpoints in the MODEL sheet (see below), by multiplying the number of cases (cell S334-352) with $DALY_{case}$ and COI_{case}, respectively.

Endpoint of the MODEL sheet

The MODEL sheet ends with the endpoints of the sQMRA model, subdivided in individual and aggregated portion categories. This includes at portion level the probability of illness and the variability therein, and the percentage of portions that leads to a human case. Further, at population level the number of human cases, the number of DALYs and the C.O.I. for every portion category are given (see Figure 9, note that the intermediate calculation results in columns J:R are not shown).

	E	F	G	H	I	S	T	U
312	**AGGREGATED PORTION CATEGORIES**							
313	**portion category**					**cases**	**DALY**	**C.O.I.**
314						number		euro
315								
316		room				8472	8,47	8,5E+05
317	+	fridge				7,9E+04	79	7,9E+06
318		freezer				7053	7,05	7,1E+05
319								
320	+	cc				6,1E+04	61	6,1E+06
321		no cc				3,4E+04	34	3,4E+06
322								
323		done				3,0E+04	30	3,0E+06
324	+	undercooked				3,9E+04	39	3,9E+06
325		raw				2,6E+04	26	2,6E+06
326								
327	-					0	0	0
328								
329								
330								
331	**INDIVIDUAL PORTION CATEGORIES**							
332						**cases**	**DALY**	**C.O.I.**
333	retail	storage			preparation	number		curr. W
334					done	3220	3	3,2E+05
335		room			undercooked	1668	2	1,7E+05
336					raw	932	1	9,3E+04
337					done	2,3E+04	23	2,3E+06
338	+	fridge	cc		undercooked	1,7E+04	17	1,7E+06
339					raw	1,1E+04	11	1,1E+06
340					done	1403	1	1,4E+05
341		freezer			undercooked	1556	2	1,6E+05
342					raw	1214	1	1,2E+05
343					done	214	0	2,1E+04
344		room			undercooked	1498	1	1,5E+05
345					raw	940	1	9,4E+04
346					done	1834	2	1,8E+05
347	+	fridge	no cc		undercooked	1,6E+04	16	1,6E+06
348					raw	1,1E+04	11	1,1E+06
349					done	163	0	1,6E+04
350		freezer			undercooked	1485	1	1,5E+05
351					raw	1232	1	1,2E+05
352	-					0	0	0
353						9,5E+04	95	9,5E+06

Fig.9: Endpoint of modelsheet

4 RESULTS sheet

4.1 Structure

On the RESULTS sheet, input parameters and output (attribution of exposure and effect, for storage and preparation categories; relative risk on portion and population level; variability on portion level; statistical uncertainty in food data) are presented. See Fig. 10 below for an overview.

Fig. 10: Overview of RESULTS sheet

4.2 Symbols

On the RESULTS sheet, calculations on the population level (e.g. attribution of exposure and effect and relative risk) are executed. To describe these calculations, a set of symbols is used which is given in the table below.

Symbol	Description
N	Number of portions
f^+	Fraction of contaminated portions
N^+	Number of contaminated portions
D	Total number of cfu for a category
$\overline{d}$	Mean number of cfu per portion
$\overline{p}$	Mean probability
I	Number of a specific health outcome (infections, illness, DALY, C.O.I.)

The subscript of a symbol can describe the point or transmission route of the risk assessment: *cont* for contamination, *r* for retail, *st* for storage or *stx* for storage transmission route *x*, *pr* for preparation, *cc* for cross-contamination (transmission route), *he* for heating or *hez* for heating transmission route *z*, *inf* for infection, *ill* for illness, *DALY* for disability adjusted life years and *COI* for cost of illness. A '+' superscript indicates actual contamination and a line above the symbol indicates the mean value.
The possible second part of the subscript, separated from the first part by a '/', describes portion categories:

- contamination at retail (+) or not (-);
- storage location *stx* : room (*r*), fridge (*f*) or freezer (*fre*);
- possible occurrence of cross contamination *ccy*: yes (*cc*+) or no (*cc*-);
- preparation method *prz*: done (*d*), undercooked (*u*) or raw (*r*).

For example $N_{pr/+,str,cc-,prd}$ stands for the number of portions after preparation that were contaminated at retail, were stored at room temperature, did not cause cross contamination and were prepared done. When a symbol represents all categories, the second part of the subscript is left out for the readers' convenience.

4.3 Input parameters

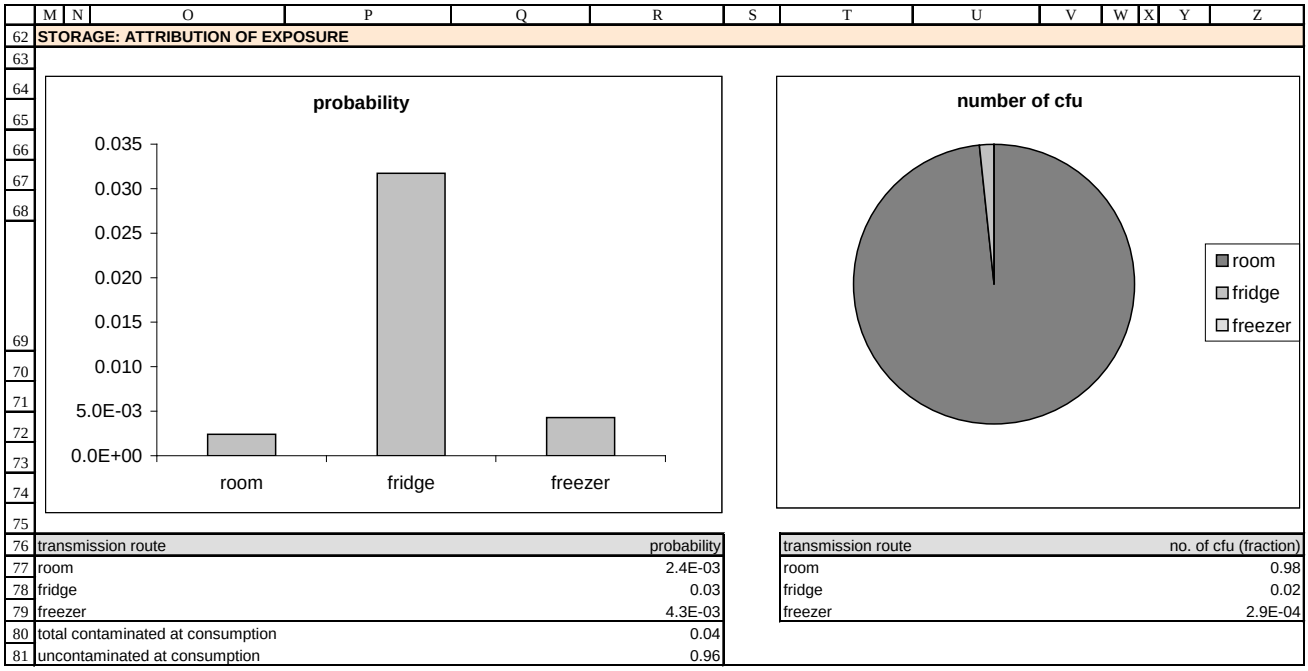
The first section of the RESULTS sheet (named 'input parameters'), starts with repeating the scope, see Ch. 3.3.1. Then, all input parameter values from the MODEL sheet are presented in a list, including question number, shortened version of the question, parameter symbol, inserted value and unit. See Fig. 6 in Ch. 3.2.

4.4 Attribution

The attribution of exposure and effect can be examined in sQMRA2 for storage and preparation categories (cross-contamination and heating categories combined).

4.4.1 Exposure

The attribution of exposure (see Fig. 11 for storage categories as an example) is described with the probability of exposure (left in the figure) and the fraction of the number of cfu per transmission route (right).



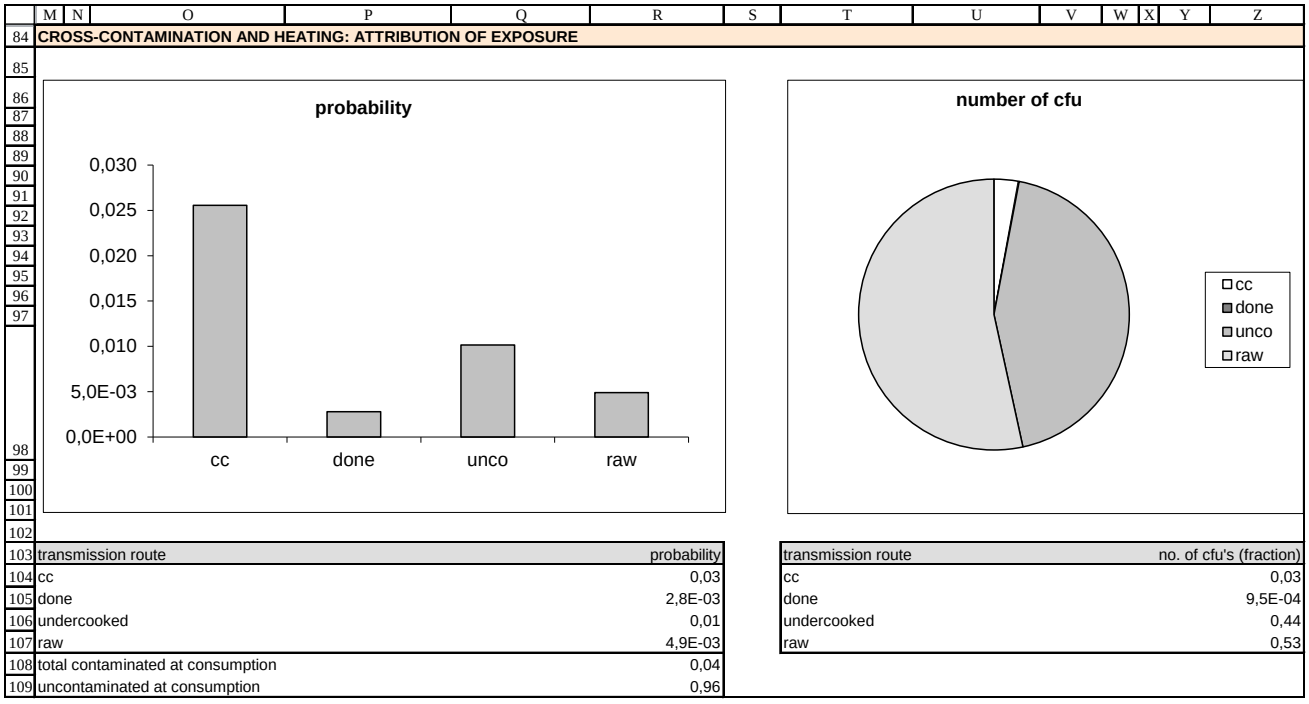
transmission route	probability
room	2.4E-03
fridge	0.03
freezer	4.3E-03
total contaminated at consumption	0.04
uncontaminated at consumption	0.96

transmission route	no. of cfu (fraction)
room	0.98
fridge	0.02
freezer	2.9E-04

Fig. 11:Storage: Attribution of exposure

Probability of exposure

The probability of exposure per transmission route is calculated by dividing the number of portions per transmission route that are contaminated at the moment of ingestion by the total number of portions. So, when ingesting a portion, the probability is $2.4 \cdot 10^{-3}$ that it was stored at room temperature and is contaminated (see cell R77 in figure 11). The overall probability of exposure, i.e. the probability that an ingested portion is contaminated, is also given, in cell R80. In case of storage categories, the sum of R77-R79 equals R80. For preparation categories, this is not so, see fig. 12.



transmission route	probability
cc	0,03
done	2,8E-03
undercooked	0,01
raw	4,9E-03
total contaminated at consumption	0,04
uncontaminated at consumption	0,96

transmission route	no. of cfu's (fraction)
cc	0,03
done	9,5E-04
undercooked	0,44
raw	0,53

Fig. 12: Cross contamination & heating: Attribution of exposure

This is due to the fact that the preparation categories are not exclusive: portions that are positive due to cross-contamination are part of the probability of exposure via cross-contamination regardless whether they are positive due to incomplete inactivation by heating or not, and vice versa. So the probabilities of exposure due to cross-contamination and incomplete heating simultaneously are counted double, and the sum of the probabilities of exposure via cc, done, undercooked and raw (cell R104-107) is larger than the total probability of exposure at consumption (R108) in which this double counting does not occur. We chose this way of presentation in order to evade presenting 9 probability of exposure categories.

Number of cfu

The fraction of the number of ingested cfu per transmission route is calculated by dividing the number of ingested cfu per transmission route by the total number of ingested cfu. So, from figure 11 we can e.g. read that 98% (cell Z77) of all ingested cfu originated from a food product which was stored at room temperature. As opposed to the previous paragraph, these fractions add up to one both for storage and for preparation categories and can therefore be presented as a pie chart.

4.4.2 Effect

An example of the attribution of effect is shown in figure 13 for preparation categories. For the attribution of effect, which is expressed as a fraction of the total number of cases, a general concept is used which is analogous to the Population Attributable Risk (PAR) in public health epidemiology studies: we estimate the relative decrease in number of cases when a transmission route is switched off. In formula:

attribution of cases = $\frac{I_{ill/base} - I_{ill/attr}}{I_{ill/base}}$

where $I_{ill/base}$ is the number of cases in the base scenario and $I_{ill/attr}$ is the number of cases when a transmission route is switched off.

Note that in general, due to this PAR-like approach, fractions do not add up to 1. In case of storage categories, the fractions do always add up to 1 due to the fact that these categories are fully separated. This implies also that a simpler formula can be used in this case: dividing the number of cases due to a transmission route by the total number of cases. Figure 13 shows that switching off the cross-contamination gives a 31% reduction of the number of cases. Note that in general (and in case of preparation categories, see Fig. 13), due to this PAR-like approach fractions do not add up to 1. In case of storage categories, the fractions do always add up to 1 because these categories are fully separated. This implies also that a simpler formula can be used in this case: dividing the number of cases due to a transmission route by the total number of cases.

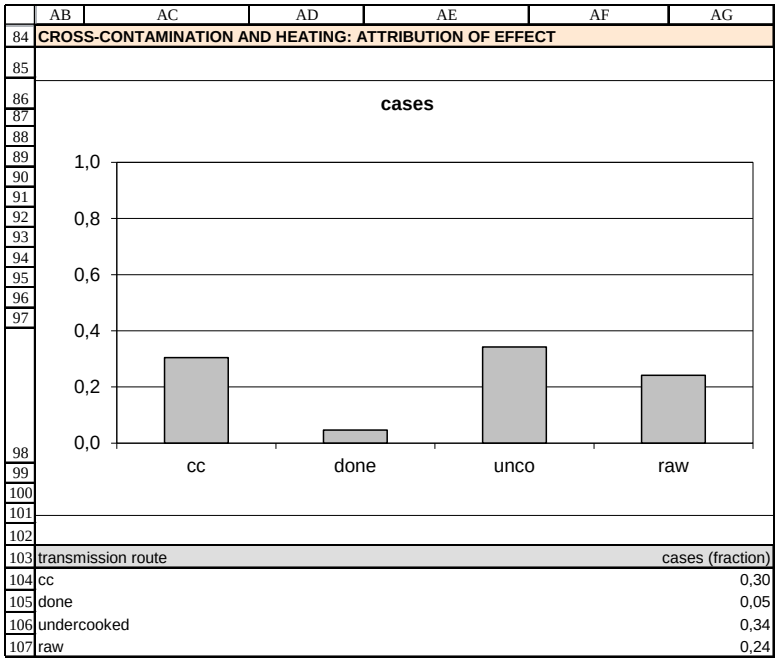


Fig. 13: Preparation: attribution of effect

4.4.3 Formulas

Every attribution value is calculated with a formula as shown in the tables below. We give the formulas in the same order as implemented in sQMRA2. The numerator in most formulas looks complex, but on many occasions simply points to one Excel cell.

Table 1: Storage: probability of contamination of a portion

Cell	Transmission route	Formula
R77	Room	$\frac{\sum_{y,z} N_{pr/+,str,ccy,prz} f^{+}_{pr/+,str,ccy,prz}}{N}$
R78	Fridge	$\frac{\sum_{y,z} N_{pr/+,stf,ccy,prz} f^{+}_{pr/+,stf,ccy,prz}}{N}$
R79	Freezer	$\frac{\sum_{y,z} N_{pr/+,stfre,ccy,prz} f^{+}_{pr/+,stfre,ccy,prz}}{N}$
R80	total contaminated at consumption	$\frac{\sum_{x,y,z} N_{pr/+,stx,ccy,prz} f^{+}_{pr/+,stx,ccy,prz}}{N}$
R81	Uncontaminated at consumption	$1-\frac{\sum_{x,y,z} N_{pr/+,stx,ccy,prz} f^{+}_{pr/+,stx,ccy,prz}}{N}$

Table 2: Storage: fraction of the number of ingested cfu at the population level

Cell	Transmission route	Formula
Z77	Room	$\frac{\sum_{y,z} D_{pr/+,str,ccy,prz}}{D_{pr}}$
Z78	Fridge	$\frac{\sum_{y,z} D_{pr/+,stf,ccy,prz}}{D_{pr}}$
Z79	Freezer	$\frac{\sum_{y,z} D_{pr/+,stfre,ccy,prz}}{D_{pr}}$

Table 3: Storage: fraction of the total number of cases

Cell	Transmission route	Formula
AG77	Room	$\frac{\sum_{y,z} I_{ill/+,str,ccy,prz}}{I_{ill}}$
AG78	Fridge	$\frac{\sum_{y,z} I_{ill/+,stf,ccy,prz}}{I_{ill}}$
AG79	Freezer	$\frac{\sum_{y,z} I_{ill/+,stfre,ccy,prz}}{I_{ill}}$

Table 4: Cross-contamination & heating: probability of contamination of a portion

Cell	Transmission route	Formula
R104	Cc	$\frac{\sum_{x,z} N_{pr/+,stx,cc+,prz} f^{+}_{cc/+,stx,cc+,prz}}{N}$
R105	Done	$\frac{\sum_{x,y} N_{pr/+,stx,ccy,prd} f^{+}_{he/+,stx,ccy,prd}}{N}$
R106	undercooked	$\frac{\sum_{x,y} N_{pr/+,stx,ccy,pru} f^{+}_{he/+,stx,ccy,pru}}{N}$
R107	Raw	$\frac{\sum_{x,y} N_{pr/+,stx,ccy,prr} f^{+}_{he/+,stx,ccy,prr}}{N}$
R108	total contaminated at consumption	$\frac{\sum_{x,y,z} N_{pr/+,stx,ccy,prz} f^{+}_{pr/+,stx,ccy,prz}}{N}$
R109	uncontaminated at consumption	$1-\frac{\sum_{x,y,z} N_{pr/+,stx,ccy,prz} f^{+}_{pr/+,stx,ccy,prz}}{N}$

Table 5: Cross-contamination & heating: fraction of the number of ingested cfu at the population level

Cell	Transmission route	Formula
Z104	Cc	$\frac{\sum_{x,z} Dei_{pr/+,stx,cc+,prz}}{D_{pr}}$
Z105	Done	$\frac{\sum_{x,y} Dhe_{pr/+,stx,ccy,prd}}{D_{pr}}$
Z106	Undercooked	$\frac{\sum_{x,y} Dhe_{pr/+,stx,ccy,pru}}{D_{pr}}$
Z107	Raw	$\frac{\sum_{x,y} Dhe_{pr/+,stx,ccy,prr}}{D_{pr}}$

Table 6: Cross-contamination & heating: fraction of the total number of cases

Cell	Transmission route	Formula
AG104	Cc	$1-\frac{\sum_{x,z} I_{ill/+,stx,cc-,prz}}{(1-S_{cc+})I_{ill}}$
AG105	Done	$1-\frac{\sum_{x,y,q\neq d} I_{ill/+,stx,ccy,prq}+\sum_{x} I_{ill\_cc/+,stx,cc+,prd}}{I_{ill}}$
AG106	Undercooked	$1-\frac{\sum_{x,y,q\neq u} I_{ill/+,stx,ccy,prq}+\sum_{x} I_{ill\_cc/+,stx,cc+,pru}}{I_{ill}}$
AG107	raw	$1-\frac{\sum_{x,y,q\neq r} I_{ill/+,stx,ccy,prq}+\sum_{x} I_{ill\_cc/+,stx,cc+,prr}}{I_{ill}}$

4.5 Relative risk

To put the model outputs in perspective, intermediate and end-point model outputs can be compared with data from reference studies (see Fig. 15 in Ch..5 for some examples).

By using the drop-down list in cells S113:Y113 in Fig. 14, a relevant study can be selected. sQMRA2 automatically shows relative risk values on the portion level and on the population level.

4.5.1. Portion level

	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
113	**RELATIVE RISK REFERENCE:**							sQMRA: Campylobacter - Chicken fillet						
114														
115	**RELATIVE RISK: PORTION LEVEL**													
116	**CONTAMINATED PORTIONS**							**ALL PORTIONS**						
117	point of comparison			model	reference	relative risk		model	reference					relative risk
118	retail (prob. of cont.)			1,00	1,00	1,00		0,10	0,40					0,25
119	storage (prob. of cont.)			1,00	1,00	1,00		0,10	0,19					0,51
120	preparation (prob. of cont.)			1,00	1,00	1,00		0,04	1,2E-03					31,91
121	retail (mean no. of cfu)			1,4E+04	2708	5,17		1,4E+03	1,1E+03					1,29
122	storage (mean no. of cfu)			1,5E+06	1174	1,3E+03		1,4E+05	225,00					640,78
123	preparation (mean no. of cfu)			3,4E+05	38	9,0E+03		1,3E+04	0,05					2,9E+05
124	prob. of infection			0,02	0,07	0,30		8,1E-04	8,4E-05					9,69
125	prob. of illness			2,1E-03	0,02	0,09		8,1E-05	2,8E-05					2,95
126	daly per portion			2,1E-06	9,3E-04	2,3E-03		8,1E-08	1,1E-06					0,07
127	c.o.i. per portion (euro)			0,21	7,95	0,03		8,1E-03	9,6E-03					0,85

Fig. 14: Relative risk reference & relative risk at portion level.

Within the relative risk on the portion level we show output data for all *contaminated* portions (Fig. 14, cells P118-R127) and the relative risk for *all* portions (cell, T118-Z127). In the first column, the model output for several 'points of comparison' is given, then the model output from the selected reference study is presented, and finally the relative risk (model output divided by reference output) is calculated. From Fig. 14 we read for instance that the probability of infection from a contaminated portion is more than two times (0.39) times lower in the example sQMRA then in the default reference study. In table 7, we show the formulas for calculating the model output for the different points of comparison.

Table 7: Relative risk on portion level. (Cell references are included between parentheses)

Point of comparison	Contaminated portions	All portions
retail (prob. of cont.)	1 (P118)	$\frac{N_r^+}{N}$ (T118)
storage (prob. of cont.)	1 (P119)	$\frac{N_{st}^+}{N}$ (T119)
ingestion (prob. of cont.)	1 (P120)	$\frac{N_{pr}^+}{N}$ (T120)
retail (mean no. of cfu)	$\frac{D_r}{N_r^+}$ (P121)	$\frac{D_r}{N}$ (T121)
storage (mean no. of cfu)	$\frac{D_{st}}{N_{st}^+}$ (P122)	$\frac{D_{st}}{N}$ (T122)
ingestion (mean no. of cfu)	$\frac{D_{pr}}{N_{pr}^+}$ (P123)	$\frac{D_{pr}}{N}$ (T123)

prob. of infection	$\frac{I_{\inf}}{N_{pr}^{+}}$ (P124)	$\frac{I_{\inf}}{N}$ (T124)
prob. of illness	$\frac{I_{ill}}{N_{pr}^{+}}$ (P125)	$\frac{I_{ill}}{N}$ (T125)
daly per portion	$\frac{I_{DALY}}{N_{pr}^{+}}$ (P126)	$\frac{I_{DALY}}{N}$ (T126)
c.o.i. per portion (euro)	$\frac{I_{COI}}{N_{pr}^{+}}$ (P127)	$\frac{I_{COI}}{N}$ (T127)

4.5.2 Population level

	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
129	**RELATIVE RISK: POPULATION LEVEL**													
130	**MODELOUTPUT**							**PER 100.000 PERSONS**						
131	point of comparison					model		model	reference					relative risk
132	portions consumed					1,2E+09		1,2E+06	2,4E+06					0,51
133	retail (no. of contaminated portions)					1,2E+08		1,2E+05	9,6E+05					0,13
134	storage (no. of contaminated portions)					1,2E+08		1,2E+05	4,6E+05					0,26
135	preparation (no. of contaminated portions)					4,7E+07		4,7E+04	2893					16,18
136	retail (no. of cfu)					1,7E+12		1,7E+09	2,6E+09					0,65
137	storage (no. of cfu)					1,8E+14		1,8E+11	5,4E+08					324,84
138	preparation (no. of cfu)					1,6E+13		1,6E+10	1,1E+05					1,5E+05
139	number of infections					9,9E+05		987	201					4,91
140	number of cases					9,9E+04		99	66					1,50
141	population DALY					99		0,1	2,7					0,04
142	population C.O.I. (euro)					9,9E+06		9871	2,3E+04					0,43

Fig. 15: Relative risk reference & relative risk at population level.

Within the relative risk on the population level, we show in the first column the model output data for the 'points of comparison' (cells R132-R142), see table 8 for formulas. We then recalculate the model output to a standardized output per 100.000 persons, to make comparisons with international studies easier (cells T132-T142). Finally, we divide the standardized model output by a reference study output (also expressed per 100.000 persons) to arrive at the relative risk on population level (cells Z132-Z142).

Table 8: Relative risk on population level

Cell	**Point of comparison**	Formula
R132	*portions consumed*	N
R133	*retail (no. of contaminated portions)*	$N_r^+ = N_{r/+} f_{r/+}^+$ *
R134	*storage (no. of contaminated portions)*	$N_{st}^+ = \sum_x N_{st/+,stx} f_{st/+,stx}^+$
R135	*preparation (no. of contaminated portions)*	$N_{pr}^+ = \sum_{x,y,z} N_{pr/+,stx,ccy,prz} f_{pr/+,stx,ccy,prz}^+$
R136	*retail (no. of cfu)*	$D_r = N_{r/+} \overline{d}_{r/+}$

R137	*storage (no. of cfu)*	$D_{st} = \sum_x N_{st/+,stx} \bar{d}_{st/+,stx}$
R138	*preparation (no. of cfu)*	$D_{pr} = \sum_{x,y,z} N_{pr/+,stx,ccy,prz} \bar{d}_{pr/+,stx,ccy,prz}$
R139	*number of infections*	$I_{\inf} = \sum_{x,y,z} N_{pr/+,stx,ccy,prz} \bar{p}_{\inf/+,stx,ccy,prz}$
R140	*number of cases*	$I_{ill} = I_{\inf} p_{ill\|\inf}$
R141	*population DALY*	$I_{DALY} = I_{ill} DALY_{case}$
R142	*population C.O.I. (euro)*	$I_{COI} = I_{ill} COI_{case}$

* N_r^+ is equal to $N_{r/+}$ minus the 'poisson-zero's' (see Ch. 6, section 'poisson effect at low concentrations and small portion size')

4.6 Other sQMRA2 output

4.6.1 Variability on portion level

	AB	AC	AD	AE	AF	AG
115	**VARIABILITY ON PORTION LEVEL**					
116	**ALL PORTIONS**					
117	point of comparison		non-zero's	variability non-zero's		
118			fraction	2.50%	mean	97.50%
119	retail (no. of cfu)		0.10	11	1.4E+04	8.9E+04
120	storage (no. of cfu)		0.10	1.9E+04	2.8E+05	1.1E+06
121	preparation (no. of cfu)		0.04	2350	5.3E+04	4.8E+05
122	prob. of infection		0.04	5.1E-05	0.02	1.00
123	prob. of illness		0.04	5.1E-06	2.1E-03	0.10

Fig. 16: Variability on portion level: fraction non-zero's and variability of non-zero's.

On the MODEL sheet we consistently show data on variability, but always divided into portion categories. In the Relative Risk section of the RESULTS sheet we show overall results (all categories combined), but give little information on variability. In the variability on portion level section (Fig. 16), we show overall results on variability at the following points of comparison: number of cfu after retail, storage, and preparation and effect measures (probability of infection and illness).

As for the fraction of non-zero portions, cells AD119-121 are equal to the probability of contamination for all portions, *pcont,f* in cells T118-T120 (see Fig. 14). Portions that contain a non-zero number of pathogens after preparation give a non-zero probability of infection *pinf* and illness *pill* and therefore the value of cells AD122-123 is equal to AD121. In case mean of non-zero portions is chosen as a measure of central tendency, cells AF119-121 are equal to the mean number of cfu on a contaminated portion , $\bar{d}_f^+$, in cells P121-123 and

cells AF122-123 are equal to the mean probability of infection or illness after consuming a contaminated portion, $\overline{p}_g^+$, in cells P124-125 (see Fig. 14).

As for the lower percentile values of the non-zero pinf and pill values (cel AE122-123), please note that numerical limitations of Microsoft Excel can result in values of zero, which however must be interpreted as very small values.

In order to increase understanding of what is presented here, in Table 9 we give the formulas for the calculations of the means of non-zero's (no. of cfu per portion or probabilities of infection or illness) as an example.

Table 9: Variability on portion level: calculation of the mean of non-zero's.

Cell	Point of comparison	Formula for mean of non-zero's
AF119	retail (no. of cfu)	$\frac{D_r}{N_r^+}$
AF120	storage (no. of cfu)	$\frac{D_{st}}{N_{st}^+}$
AF121	preparation (no. of cfu)	$\frac{D_{pr}}{N_{pr}^+}$
AF122	prob. of infection	$\frac{I_{\inf}}{N_{pr}^+}$
AF123	prob. of illness	$\frac{I_{ill}}{N_{pr}^+}$

4.6.2 Statistical uncertainty in retail data

Many parameters in sQMRA2 have the option to describe variability by means of an appropriate probability distribution. Uncertainty is not taken into account in the model calculations. We include an example of uncertainty to confront the sQMRA user with the important difference between variability and uncertainty. For questions 1 and 2 ('portions consumed' and 'pathogen prevalence at retail') statistical uncertainty therefore, can easily be calculated through means of the appropriate probability distributions and is presented in cells AB125-AG131 (see figure below). The used probability distributions are given in table 10.

	AB	AC	AD	AE	AF	AG
125	**STATISTICAL UNCERTAINTY IN RETAIL DATA**					
126						
127			iteration	uncertainty		
128				2,50%	mean	97,50%
129	no. of portions		1,8E+09	5,8E+08	1,2E+09	2,1E+09
130	prevalence in retail		0,09	0,08	0,10	0,12
131	no. of cont. portions		1,7E+08	5,7E+07	1,2E+08	2,1E+08

Fig. 17: Statistical uncertainty of consumption and prevalence data.

Table 10: Statistical uncertainty in retail data

Cells	**Retail data**	Formula
AD129	no. of portions	Gamma[N_{survey}, $1/t_{survey}$]*Pop*t_{cons}
AD130	prevalence in retail	Beta[Pos + 1, $Size$ - Pos + 1]
AD131	no. of cont. Portions	Cell AD129 * Cell AD130

4.6.3 sQMRA2: simulation info

In cells AB135-AG138 (see figure below), information on the current filename, current date and time (for printing purposes) and total number of iterations in the current simulation are provided. In cells AB139-AG142 contact information and the logos of the collaborating institutes are presented.

	AB	AC	AD	AE	AF	AG
133	**sQMRAv2: simulation info**					
134						
135	this simulation result:					
136	filename				sQMRA v2_95 - tbv handleiding2.xls	
137	current date and time				7-aug-2012	12:35
138	total no. of iterations:					10000
139						
140						
141						
142	contact: eric.evers@rivm.nl, jurgen.chardon@rivm.nl					

Fig. 18: Simulation file information.

5 REFERENCE DATA sheet

On the REFERENCE DATA sheet, the user can enter pathogen-product data or calculation results in the blue data entry cells. It is important to note that reference data has to represent a population of 100.000 persons. The table in cells C2:P18 (partly shown in figure 19) consists of 5 user defined and 7 predefined pathogen-product combinations. Once a reference pathogen-product combination is entered, the combination can be selected with the drop down list in cells S113:Z113 on the RESULTS sheet, see also Fig. 14.

	C	D	E	F	G	H	I	J
2	**REFERENCE ADMINISTRATION**							
3	**INSTRUCTIONS**: REFERENCE DATA NEED TO BE ENTERED AS NUMBERS PER 100.000 persons							
4								
5		**SELECTED REFERENCE**						
6	**Points of comparison**	**sQMRA: Campylobacter - Chicken fillet**	**user defined 1**	**user defined 2**	**user defined 3**	**sQMRA: Campylobacter - Chicken fillet**	**sQMRA: Salmonella - Chicken fillet**	**sQMRA: Campylobacter - Fillet Americain**
7	portions consumed	**2,4E+06**	0	0	0	2,4E+06	2,4E+06	4,3E+05
8	cont. Portions | retail	**9,6E+05**	0	0	0	9,6E+05	2,0E+05	1574
9	cont. Portions | storage	**4,6E+05**	0	0	0	4,6E+05	1,3E+05	1573
10	cont. Portions | preparation	**2893**	0	0	0	2893	1686	1573
11	total number of cfu | retail	**2,6E+09**	0	0	0	2,6E+09	8,0E+07	8,1E+05
12	total number of cfu | storage	**5,4E+08**	0	0	0	5,4E+08	3,5E+11	5,0E+05
13	total number of cfu | preparatio	**1,1E+05**	0	0	0	1,1E+05	5,3E+08	5,0E+05
14	number of infections	**201**	0	0	0	201	46	568
15	cases of illness	**66**	0	0	0	66	46	188
16	dalys	**3**	0	0	0	3	2	8
17	cost-of illness	**2,3E+04**	0	0	0	2,3E+04	1,1E+04	6,5E+04
18								

Fig. 19: Reference data.

In cells C36:Q47 (not shown in Fig. 19) the ‘references for relative risk reference data’ are presented. Extended references are available upon request.

6 sQMRA2 calculation considerations

In this chapter we present some considerations with respect to the execution of sQMRA2 calculations. We give information about:

- Variability and uncertainty
- True and measured prevalence and concentration
- Poisson effect at low concentrations and small portion size
- Calculation of statistics
- Using RiskDiscrete for aggregated portion categories
- @RISK output graphs
- Convergence

Variability and uncertainty

"Variability is the effect of chance and is a function of the system. It is not reducible through either study or further measurement. (…) Uncertainty is the assessor's lack of knowledge (level of ignorance) about the parameters that characterise the physical system that is being modelled" (Vose, 2008)

In sQMRA2, variability can be included optionally through means of probability distributions. Uncertainty is not included in sQMRA2. If one wishes to study uncertainty of parameter values, we suggest to run simulations with different parameter values as a scenario analysis.

The default set of parameter values in sQMRA2 is deterministic. To give the user some guidance on the importance of including variability for the different sQMRA2 parameters in relation to model output, an analysis of variability has been performed for the pathogen-product combinations Campylobacter *spp* in Chicken, Salmonella *spp.* in Table eggs and Listeria *monocytogenes* in Filet americain. We studied the impact of including variability for estimation of the number of cases (population level) as well as for estimation of the variability between portions (portion level). From the results of this analysis, we give the following general recommendations:

Variability relevant for estimation of the number of cases and the variability between portions:

- Storage temperature in fridge → Enter data for mean *and* standard deviation if available.
- Storage time in fridge → Enter data for mean *and* maximum storage time if available.
- Dose response model → Choose the beta-binomial dose response model if possible.

Variability relevant only for estimation of the variability between portions:

- Concentration of cfu → Enter data for mean *and* standard deviation if available.
- Cross-contamination → Enter data for mean *and* standard deviation if available.

True and measured prevalence and concentration

It is important to note that the prevalence and concentration that are requested in the tool in questions 2 and 4 are the true values, which can only be approximated by -and are usually higher than- the measured values. This should as far and as much as possible be taken into account when inserting values for these questions. When looking at quantitative measurements, the applied method can result in lower estimated concentrations than the true values (recovery < 100%). If there is a detection limit, presence will not be demonstrated in samples containing lower pathogen concentrations, and this may result in a lower estimated prevalence than the true value. Beside this, low concentrations (numbers actually) can result in absence of the pathogen in a sample by coincidence (the so-called Poisson zeros in case of random distributed micro-organisms; see next section), which also results in a lower estimated prevalence than the true value.
Also when the prevalence is determined as absence/presence through enrichment, the applied method and coincidence can result in a lower estimated prevalence than the true value.
Overall, the aspects in this section can be summarized with the term sensitivity, which may be defined as the probability to detect the pathogen when it is present in the product.

Poisson effect at low concentrations and small portion size.

With low concentrations of cfu and a small portion size, a fraction of the simulated retail-contaminated portions will have 0 cfu in the sQMRA2 model. This is comprehensible if we bear in mind that the sampling of contaminated portions is a Poisson process.

For a portion of 10 grams with a mean concentration of -1.0 $\log_{10}$ cfu/g (= 0.1 cfu/g), 36,8% of the contaminated portions will have 0 cfu. When we consider a portion of 50 grams with -1.0 $\log_{10}$ cfu/g (= 0.1 cfu/g), the ‘Poisson effect’ generates a percentage zero’s of only 0.7%.

@RISK output graphs

The standard @RISK graphing functionality can be used in sQMRA2. Simply select the yellow iteration cell for which you want to get a graph and add an output (Excel menu-item: @RISK/Model/Add Output). Then run the simulation and browse the results (Excel menu-item: @RISK/Results/Browse Results). Note that @RISK has no option to plot data on the logarithmic scale.

Calculation of statistics

Throughout the tool, output variability is presented (mainly % non-zero’s, percentiles, mean, median and mode). Standard @RISK and Excel functions are used to calculate this output. Here, we present the general technical set-up of this output. This setup applies to both the individual portion categories as well as to the aggregated portion categories.

% non-zero's

The percentage of non-zero's ($percent_{nz}$) produced in yellow iteration cells (y_{ellow}) is calculated with '=(1-RiskTarget(y_{ellow},0)', where RiskTarget returns the cumulative probability of 0 for y_{ellow}.

Mean

Knowing the percentage of non-zero's (see above), the following formula applies to calculate the mean for the non-zero iteration values: '=RiskMean(y_{ellow} / p_{nz})

Percentiles and median

Knowing the percentage of non-zero's (see above), the following formula applies to calculate user-defined percentiles ($percent_{user}$) for the non-zero iteration values: '=RiskPercentile(y_{ellow}, 1- $percent_{nz}$ + $percent_{user}$*$percent_{nz}$).

Mode

Finally, for the mode we use the standard @RISK RiskMode-function. Note that we always use RiskMode in combination with the RiskTruncate-function to filter out the zero's generated in the yellow iteration cells. For the application of the RiskTruncate-function, please refer to your @RISK software documentation.

Using RiskDiscrete for aggregated portion categories

In sQMRA2 we use the @RISK RiskDiscrete-function to aggregate portion categories. For instance, if room, fridge and freezer are categories with 2, 3 and 5 portions and mean storage times of 1, 2 and 9 days, then the aggregated mean storage time, regardless of storage location, is (2*1 + 3*2 + 5*9)/10 = 5.3 days. We would implement this in @RISK by reading out the RiskMean value for '=RiskDiscrete({1,2,9},{2,3,5})'. See also your @RISK software documentation.

Convergence

We executed a series of simulations with the pathogen-product combinations 'Salmonella in Eggs' and 'Campylobacter in Chicken' to determine convergence with bandwidths of 1%, 3% and 5% for the output 'number of illness'. The user can use these values as a guide for the appropriate number of iterations needed for sQMRA2 simulations.

Pathogen-product combination	Bandwith		
	1%	3%	5%
Salmonella in Eggs	20.000	2500	2500
Campylobacter in Chicken	50.000	20.000	10.000

7 Technical considerations

In this chapter, some technical considerations are given about default sQMRA2 settings:

- Input validation
- Sheet protection
- Decimal separator
- @RISK simulation settings
- Approximating the @RISK RiskBinomial function
- Excel background error checking
- Excel security settings

Input validation

Input validation is implemented at a minimal level to allow maximum flexibility when experimenting with parameter values in sQMRA2:

- Because of numerical limitations of the @RISK RiskGamma function, the standard deviation for the size of a portion (question 3) has to be smaller than 10 times the mean size. This is implemented as: N50 = (IF(M50/L50<10;"";"The coefficient of variation (st.dev / mean) > 10: Please adjust st. dev!")).
- For both the storage and the preparation categories (questions 5 and 8), the fractions have to add up to 100%. This is implemented as: N78 = IF(K78+L78+M78=1;"";"Fractions must add up to 100%") for storage and N175 = IF(K175+L175+M175=1;"";"Fractions must add up to 100%") for preparation.

Sheet protection

By default, the MODEL, RESULTS and REFERENCE DATA sheets are write protected except for the blue data entry cells. This protection is implemented to minimize the occurrence of unwanted modification or deletion of Excel formulas. If a user wants to modify the design of sQMRA2 to fit better to a certain risk question, the protection can be disabled easily by selecting Excel menu-item: Tools/Protection/Unprotect Sheet. To guarantee that the disabling of sheet protection is a conscious action, the Excel sheet protection is restored automatically upon reopening of the file.

Decimal separator

In sQMRA2 it is advisable to use a dot (“.”) as the decimal separator. If necessary, configure these settings in your Excel software.

@RISK simulation settings

When opening a new sQMRA2 Excel file, the following @RISK simulation settings are default (Excel menu-items: @RISK/Settings/Simulation settings/….):

- General Tab: Number of Iterations: 1000
- View Tab: Show Excel recalculations: selected
- Sampling Tab: Update statistic functions: Each iteration

These settings are chosen for educational reasons: It gives a short simulation where statistics in the sheet change visually after each iteration. This gives a good impression of the development of a simulation.

When a higher number of iterations is important (e.g. when convergence is essential), one can deselect the 'Show Excel recalculations' item. This will speed up the simulation with a factor of 3 to 6. When in addition 'At the end of each simulation' is selected from the item 'Update statistics functions', the simulation will speed up in total with a factor of 6 to 12.

Approximating the @RISK RiskBinomial function

Due to calculation limits in MS Excel, we approximate the @RISK function RiskBinomial(*n*,*p*) with the Excel function ROUND(*n***p*,0), when *n* is equal to or bigger than 10000.

Excel background error checking

Note that for your convenience the option 'Enable background error checking' from Excel menu-item 'Tools/Options/Error Checking/Enable background error checking', is disabled. This implies that no little green triangles, which normally indicate inconsistent formulas or errors, are shown.

Excel security settings

For transparency reasons, all calculations in sQMRA2 are executed with standard @RISK worksheet functions which are directly visible in Excel cells. The only exception is VBA code used to protect worksheets. This VBA code can however generate macro security alerts. These security alerts can be disabled by decreasing the security level in the Excel menu-item 'Tools/Macro/Security/Security level…'.

References

Evers, E. G. and J. E. Chardon (2010). "A swift Quantitative Microbiological Risk Assessment (sQMRA) tool." Food Control 21(3): 319-330.

van Asselt, E. D. and M. H. Zwietering (2006). "A systematic approach to determine global thermal inactivation parameters for various food pathogens." Int J Food Microbiol 107(1): 73-82.

van Gerwen, S. J. and M. H. Zwietering (1998). "Growth and inactivation models to be used in quantitative risk assessments." J Food Prot 61(11): 1541-1549.

Vose, D. Risk Analysis: a quantitative guide, 3rd edition, 2008, John Wiley & Sons Ltd, Chichester.